Pre-print manuscript version accepted by the ISPRS Journal of Photogrammetry and Remote Sensing. CLICK for the final version.

Coregistration refinement of hyperspectral images and DSM: An object-based approach using spectral information

Janja Avbelj[a,b,*], Dorota Iwaszczuk[c], Rupert Müller[a], Peter Reinartz[a], Uwe Stilla[c]

[a] *German Aerospace Center, Oberpfaffenhofen, 82234 Weßling*
[b] *Chair of Remote Sensing Technology, Technische Universität München, Arcisstr. 21, 80333 München*
[c] *Photogrammetry and Remote Sensing, Technische Universität München, Arcisstr. 21, 80333 München*

Abstract

For image fusion in remote sensing applications the georeferencing accuracy using position, attitude, and camera calibration measurements can be insufficient. Thus, image processing techniques should be employed for precise coregistration of images. In this article a method for multimodal object-based image coregistration refinement between hyperspectral images (HSI) and digital surface models (DSM) is presented. The method is divided in three parts: object outline detection in HSI and DSM, matching, and determination of transformation parameters. The novelty of our proposed coregistration refinement method is the use of material properties and height information of urban objects from HSI and DSM, respectively. We refer to urban objects as objects which are typical in urban environments and focus on buildings by describing them with 2D outlines. Furthermore, the geometric accuracy of these detected building outlines is taken into account in the matching step and for the determination of transformation parameters. Hence, a stochastic model is introduced to compute optimal transformation parameters. The feasibility of the method is shown by testing it on two aerial HSI of different spatial and spectral resolution, and two DSM of different spatial resolution. The evaluation is carried out by comparing the accuracies of the transformations parameters to the reference parameters, determined by considering object outlines at much higher resolution, and also by computing the correctness and the quality rate of the extracted outlines before and after coregistration refinement. Results indicate that using outlines of objects instead of only line segments is advantageous for coregistration of HSI and DSM. The extraction of building outlines in comparison to the line cue extraction provides a larger amount of assigned lines between the images and is more robust to outliers, i.e. false matches.

Keywords: Hyper spectral, DEM/DTM, Registration, Multisensor, Urban, Matching, LiDAR

1. Introduction

3D city models are incorporated into many applications, e.g. urban planning, micro climate analysis, vizualisation, disaster management, indoor navigation, and many more (Rottensteiner et al., 2014; Brédif et al., 2013). The importance of the city models shows in the efforts on their standardisation. The Open Geospatial Consortium (OGC) developed and is constantly improving the City Geography Markup Language (CityGML), an international standard for representation, storage, and exchange of 3D city models (Gröger et al., 2012). Moreover, the European Union has been dealing with data specifications on buildings in the INSPIRE (Infrastructure for Spatial Information in Europe) framework (INSPIRE TWG BU, 2013). For 3D building models, both CityGML and INSPIRE framework define the 3D geometry at different levels of detail, semantics, and material attributes of building façades and roofs (Gröger and Plümer, 2012; INSPIRE TWG BU, 2013). Data from remote sensing sensors can be used to derive city models with increased performance, when evaluated in a combined framework.

Remote sensing sensors are constantly undergoing developments to improve their performance. Along with the increase in spatial resolution, new technologies are emerging, e.g. hyperspectral imaging (HSI) sensors are progressively replacing multispectral (MSI) ones. HSI sensors are able to measure reflectance in hundreds of narrow spectral bands with increased spectral resolution in comaprison to multispectral sensors, which rely on few bands (e.g. $\leq$12) of larger bandwidth, and gaps between them (Keshava, 2003). Furthermore, images acquired with various techniques have *per se* different characteristics even if observed in the same portion of the spectrum. For instance, hyperspectral data allow differentiating between materials on the basis of their rich spectral information, whereas light detection and ranging (LiDAR) measurements provide highly accurate surface models. To gain more information about the observed objects multimodal images, images acquired by different imaging techniques featuring different characteristics can be fused.

Image and data fusion are important procedures for the integration and analysis of many data sets in photogrammetry, remote sensing and computer vision. Fusion becomes no-

*Corresponding author.
Email addresses: janja.avbelj@dlr.de (Janja Avbelj), dorota.iwaszczuk@tum.de (Dorota Iwaszczuk), rupert.mueller@dlr.de (Rupert Müller), peter.reinartz@dlr.de (Peter Reinartz), stilla@tum.de. (Uwe Stilla)

tably challenging when integrating data from different sensors, mounted on different platforms, and collected at different times. Usually, the acquired data are georeferenced, but often the accuracy of the georeferencing is not sufficient to achieve the best possible fit between all the data sets. These geometric inaccuracies can lead to difficult or even impossible interpretation of the results of the fusion process. Consequently, refined coregistration of all data sets has to be carried out prior to the fusion process even though both datasets are georeferenced.

Remote sensing provides images and methods for: (a) automatic building model extraction, where building modeling methods are often based on digital surface models (DSM), derived from optical stereo images or LiDAR point clouds, and (b) information about surface materials, where HSI enable the distinction between different materials. DSM and HSI have obviously different and complementary characteristics, due to their different acquisition techniques. Thus, the accurate coregistration between HSI and DSM is a step towards the envisaged automatic extraction of roofing materials for building models using remote sensing imagery.

In the present article, we propose an original method for object-based coregistration refinement of DSM and HSI images and compare it to the line-based coregistration. Focusing on urban areas, we base the image-to-image coregistration on spatial features of buildings objects, which are described by outlines. For line-based coregistration, line segments are extracted independently of each other, whereas for object-based coregistration these are extracted using knowledge about the topology and regularities within approaches dealing with (entire) building outlines. Thus, building outlines represent an object and extend line-based approaches by incorporating topology knowledge (Avbelj, 2012). Furthermore, the geometric accuracy of the detected line segments of building outlines is taken into account. A stochastic model is introduced to compute optimal transformation parameters between the HSI and the DSM. The coregistration problem is formulated using homogeneous coordinates, as points and lines have the same representation. The method enables registration of HSI and DSM of different spectral and spatial resolution.

1.1. Related work

Coregistration of images requires detecting common objects or common features in the images. Zitova and Flusser (2003) compiled an extensive overview of registration methods, dividing them in feature- and area-based methods. Feature-based methods use extracted image cues, such as points (0D) and lines (1D) for matching, while area-based methods use image intensity values extracted from image patches (2D). Both types of image registration methods can be carried out in the image space or in the feature space. Area-based methods for multimodal image registration, using e.g. mutual information, have shown success in remote sensing and medical imaging (Suri and Reinartz, 2010; Pluim et al., 2003). Considering feature-based methods, the features detected in the multimodal data set with different characteristics do not always correspond to the objects of interest in space. Blaschke (2010) discussed the improvements of object-based over per-pixel image analysis. These are significant when the objects under consideration are much larger than the pixel size.

Buildings and other man-made objects are mainly described by linear structures. Therefore, for description and coregistration of data acquired in urban areas, many authors propose lines or line segments. Stilla (1995) used linear structures to interpret airborne images by creating an image-description graph derived from 2D maps. Habbecke and Kobbelt (2010) coregistered 2D cadastral maps with oblique aerial images in three stages using vanishing points, endpoints of vertical line segments, RANSAC (Fischler and Bolles, 1981) and a bundle-adjustment-like global optimization over all camera parameters. Iwaszczuk et al. (2012) matched thermal oblique images with 3D models for texture extraction. Other authors used linear structures for 3D reconstruction of urban areas (Debevec et al., 1996; Schmid and Zisserman, 2000; Schenk, 2004; Ok et al., 2012).

Brook et al. (2010) already proposed a method using linear features around urban objects for HSI and LiDAR registration. However, their registration is based on tie points, where the detected lines are only indirectly incorporated into image registration. In Brook and Ben-Dor (2011) the outlines of buildings and roads are used for coregistration, carried out through weighting for point-to-point matching.

The coregistration refinement of georeferenced and orthorectified HSI with a DSM can be simplified by a 2D transformation, for instance an affine transformation or its variations (e.g. 2D motion transformation). Hartley and Zisserman (2004) calculated an affine transformation from line correspondences using the direct linear transformation method (DLT). Zeng et al. (2008) presented conditioning of the line parameters in order to stabilize the solution numerically. However, they did not take into account the geometric accuracy of these features, which depends on the data and detection method used. This accuracy should also be taken into consideration for coregistration.

1.2. Overview

This article extends the coregistration refinement method for HSI and DSM in urban areas by Avbelj et al. (2013). The flowchart describing the proposed object-based coregistration method is shown in Fig. 1. The main novelty of the method is the combined use of material and height properties of urban objects for coregistration refinement. Besides, the line-to-line matching is carried out through an accumulator using statistical tests. Another novelty is the introduction of the geometric accuracy for the extracted features, which are used for the parameter transformation estimation step. Finally, the evaluation is carried out by comparing the computed inner accuracies of the transformation parameters with the reference parameters, defined by using building outlines of much higher resolution.

The structure of the article is as follows. In Section 2 we describe the material map extraction from HSI and the building outline detection from HSI and DSM. To demonstrate the importance of object-based coregistration, we also independently extract line segments from both data sets. Then, in Section 3 we propose the weighted line segment matching and transformation parameter calculation. Feasibility and applicability of the

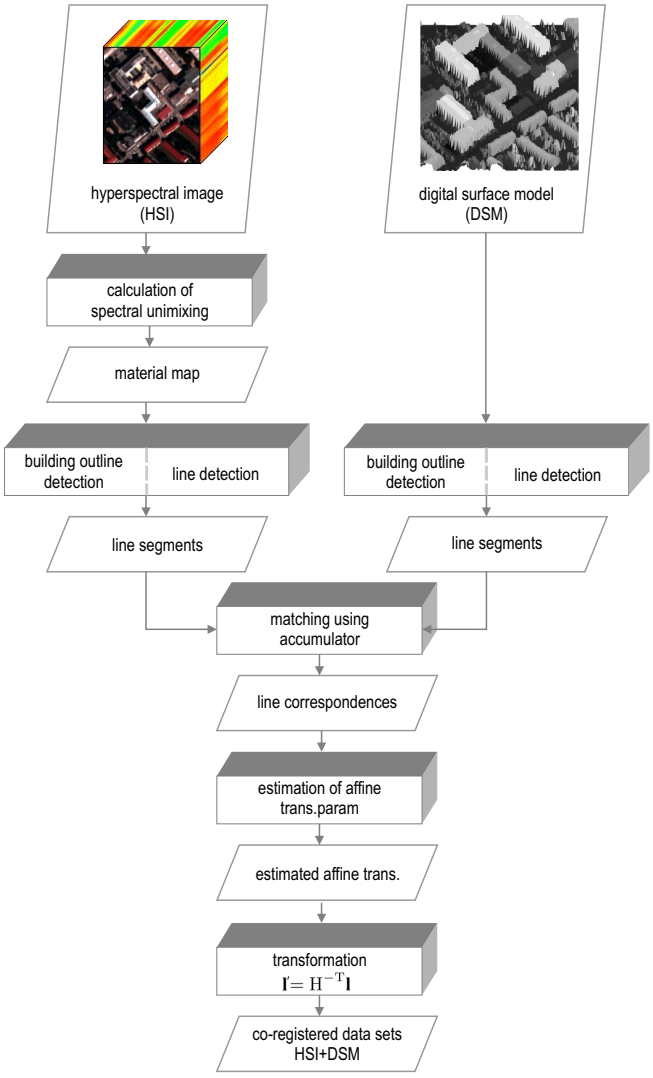


Figure 1: Workflow for theproposed object-based coregistration refinement of HSI and DSM.

proposed method is demonstrated by performing experiments on two DSM with different spatial resolution and two aerial HSI with different spatial and spectral resolution (Sec. 4). The results are presented, discussed and the accuracy of the coregistration is evaluated (Subsec. 4.3). Finally, we conclude and discuss future work in Sec. 5.

2. Building extraction and description

In this section we present the principle of building outline detection from two data sets with different characteristics, HSI and DSM. DSM are regarded as 2.5 dimensional data and represented as images, where intensity values indicate heights. The radiometric calibration, i.e. the transformation of digital numbers into a radiometric unit of measure, is a prerequisite for multi-source and multi-temporal image harmonization and interoperability (Schaepman-Strub et al., 2006; GEO/CEOSS, 2010). Skipping the radiometric calibration would negatively affect the transferability of a physical model-based image processing system. Therefore, the HSI in the present paper are first calibrated into at-sensor radiance values. Then, spectral signatures of building roofing materials are used to calculate material maps by unmixing the pixels of HSI (Subsec. 2.1). Finally, building outlines and line cues (Subsec. 2.2) are extracted from material maps and DSM, and independently used for matching and to estimate the transformation parameters as described in Section 3.

2.1. Material detection from hyperspectral images

Every material is uniquely described by a spectral signature, and can be compared to the spectra in the HSI characterized by n spectral bands. Some of the pixels in the HSI consist of more than just one material and are called mixed pixels. If a complete collection of m materials with spectral signatures or endmembers $\mathbf{S} \in \mathbb{R}^{n\times m}$ is known, then the fractions or abundances of materials $\mathbf{a} \in \mathbb{R}^m$ can be computed for every pixel $\mathbf{x} \in \mathbb{R}^n$ in the HSI (Fig. 2).

A common procedure to decompose each measured pixel into the constituent spectra is spectral unmixing. We assume a linear mixing model where: (a) the surface consists of distinct materials, (b) each surface features a Lambertian reflectance (Keshava, 2003; Schaepman-Strub et al., 2006), and (c) the number of endmembers is smaller than the number of channels: $m \leq n$. Then a pixel $\mathbf{x} \in \mathbb{R}^n$ from a HSI is expressed as

$$\mathbf{x} = \sum_{i=1}^{m} a_i \mathbf{s}_i + \mathbf{w} = \mathbf{S}\mathbf{a} + \mathbf{w} \qquad (1)$$

where a_i is the fraction of the i-th material (abundance), $\mathbf{s}_i \in \mathbb{R}^n$ is the spectral signature of i-th material (endmember), and $\mathbf{w} \in \mathbb{R}^n$ is the residual vector. The abundance vector $\mathbf{a}$ is computed by solving Eq. 1 with the least squares method. The expression for one HSI pixel (Eq. 1) can be generalized for the complete

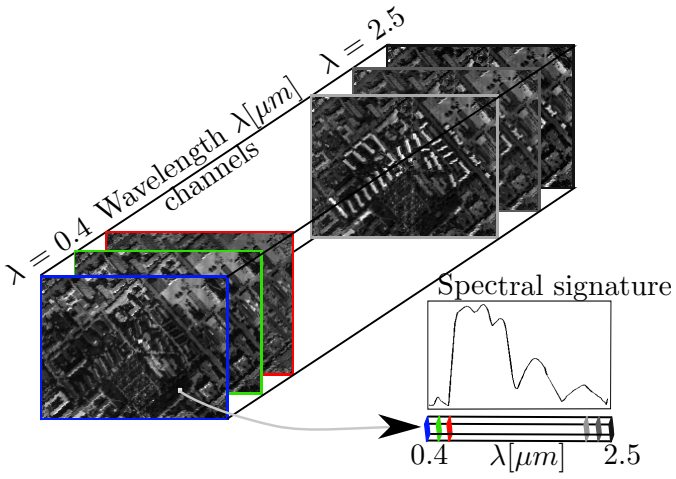


Figure 2: Spectral signature. An optical hyperspectral image consists of many spectral bands with wavelengths typically ranging from 0.4 to 2.5 μm. Each pixel in a HSI can be viewed as a function of the wavelengths and represents spectral signature.

image $\mathbf{X} \in \mathbb{R}^{n \times k}$, $\mathbf{X} = \mathbf{SA} + \mathbf{W}$, where columns of $\mathbf{X}$ are k pixels from HSI, $\mathbf{S} \in \mathbb{R}^{n \times m}$, $\mathbf{A} \in \mathbb{R}^{m \times k}$ and $\mathbf{W} \in \mathbb{R}^{n \times k}$. The solution of the HSI unmixing is a set of abundance maps, one for each material (Fig. 3).

For abundances to have a physical meaning, the linear mixing model is subject to two constraints. First, the fraction $\mathbf{a}_i$ of any material in a HSI pixel can not be negative, leading to the non-negativity constraint

$$a_i \leq 0. \tag{2}$$

Second, the sum of the abundances for any HSI pixels must be equal to one, leading to the full additivity constraint

$$\sum_{i=1}^{m} a_i = 1. \tag{3}$$

If only an incomplete collection of endmembers of the scene is available, the full additivity constrain causes over-fitting of the least squares solution. One possibility to evade the over-fitting in such cases is by using a linear mixing model subject only to the non-negativity constraint, called also non-negative least squares unmixing (NNLS). It is also noted that NNLS promotes a sparse solution for the abundance vector. Alternative possibilities are to generalize the full additivity constraint to the partial additivity constraint $\sum_{i=1}^{m} a_i \leq 1$, or to use distance metrics instead of unmixing. Distance metrics, such as spectral angle distance (Kruse et al., 1993) or spectral information divergence (Chang, 2000) compute the distance between the known spectra $\mathbf{s}_i$, $i = 1, \ldots, m$ and the image spectra $\mathbf{x}$. In contrast to unmixing algorithms, the distance metrics consider known spectra independently, so the distances are computed m times. The benefit of unmixing algorithms is that all the known spectra $\mathbf{s}_i$, $i = 1, \ldots, m$ and possible constraints are solved in a joined system. In this work we test our method by using four spectra, i.e. $m = 4$, of which three are related to roofing materials (Subsec. 4.2, Fig. 3).

To detect the buildings in HSI, spectral signatures of roofing materials present in the scene, such as metallic roofing and ceramic tiles, are manually collected from the HSI. Then, the abundances a_i of the collected materials $\mathbf{s}_i$ are computed with NNLS unmixing, resulting in m material (abundance) maps. Each of these material maps shows the spatial distribution of a roofing material and the fraction of the material. So, according to the objects of interest, the corresponding set of material maps is selected for further processing. For the unmixing process it is recommendable to use all materials present in the scene, if their spectral signatures are available. Otherwise, the unmixing should be appropriately constrained.

2.2. Building description by outlines and lines

The correspondence between a HSI and a DSM is established by detecting the same objects in both data sets. Building outlines are detected directly in the normalized DSM as above ground objects, whereas in the HSI the building outlines are detected from material maps of roofing materials (Fig. 3). To test and evaluate the importance of object-based image coregistration refinement, image cues are also extracted from the same data. So, registration between each pair of HSI and DSM is carried out twice, first by using the building object outlines, and second by using the image cues. The majority of the building outlines can be described by straight line segments, so we choose line segments as appropriate image cue. We summarize the used building outline detection in Subsec. 2.2.1, and the line cue detection in Subsec. 2.2.2, however other outline and line extraction methods could be applied.

2.2.1. Building outline detection

Many building detection and modelling methods have been proposed in the last decades, using DEM or LiDAR point clouds (e.g., Maas and Vosselman, 1999; Lafarge et al., 2008), or using additionally optical imagery (e.g., Sohn and Dowman, 2007; Awrangjeb et al., 2010). In this subsection, we summarize the building outline detection method described in Avbelj et al. (2013), but any other building outline detection could be used. The method can be applied on a HSI and a DSM and is divided in three steps: (a) building mask extraction, (b) building model selection, and (c) boundary adjustment. The minimal size for the buildings that can be extracted with the proposed method is dependent mainly on the ground sampling distance (GSD) of the image. From a mathematical point of view, this method is applicable if all sides of a building are at least three pixels long, i.e. three pixels are needed to estimate the line parameters and their accuracy. Furthermore, the sides of a building to be extracted should be at least five pixels long in order for the (b) building model selection and (c) the boundary adjustment to be reliable. The two edge pixels, in which the line segment begins and ends, are unreliable, because they span in general cases only part of the pixels.

(a) The building masks are extracted from the HSI and the DSM as an initial guess for building regions. A building mask from the HSI ($\mathbf{BM}_{HSI,i}$) is defined by thresholding a material map of every roofing material: $\mathbf{BM}_{HSI,m} > threshold$ for $i = 1, \ldots, m$. A building mask from the DSM ($\mathbf{BM}_{DSM}$) is extracted by normalizing the DSM and setting a threshold defining the above-ground objects, and removing high vegetation.

(b) The rectilinear building outlines can be reconstructed by hierarchically adding and subtracting rectangles (Gerke et al., 2001). Thus a set of building outlines is built for every region in the building mask. On each higher hierarchical level more rectangles are needed to extract a building outline. A higher cost is assigned to a more complex outline at the higher hierarchical level. Before the model of higher level is created, the boundaries are adjusted to the data as described in (c). Then, the optimal building outline model is selected as a trade off between cost and adaption to the data.

(c) The line segments of the building outline are adjusted to the image gradients using a gradient descent method. For every line segment the direction of the gradient is computed from the gradients of the pixels under the line segment. Then, the line segment is moved iteratively in the gradient direction, until a stopping criterion is satisfied, i.e. the change in the line segment position is insignificant. In our case the gradients are

(a) (b) (c) (d) (e)

Figure 3: Non-negative linear unmixing of HSI. Four image spectra were collected from the same radiometrically corrected HyMap HSI (a; Fig. 5). A HyMap HSI (subset of 70x70 pixels) is shown as RGB composite (a), and four material maps: reed roof tiles (b), metallic roof (c), concrete (d), and vegetation (e). For the material maps colour coding is from black to white, where white means the highest abundance.

computed from the material map of the corresponding roofing material and the normalized DSM. Finally, the building outline generalization is carried out by joining adjacent line segments with an intersection angle smaller than 10°.

In Section 3 the extracted outlines from both images are matched and used to refine the coregistration, more precisely the line segments that constitute the outlines are matched. Due to the building outline method described above (a)-(c), these line segments have some restrictions, e.g. the end point of a line segment must be equal to the starting point of the subsequent line segment. Furthermore, the assumption of orthogonality of the subsequent outlines‘ edges given in (b) is taken into account by a metric restriction to the building outlines. Hence, the line segments of the building outlines implicitly introduce the topology relations and some metric restrictions between subsequent edges of the building outlines into the estimation of the transformation parameters. These restrictions are not directly required for the transformation parameter estimation because all the details of the building outlines or their details might not be extractable from both of the data sets.

2.2.2. Line detection

Line segments are detected in the DSM and in the material maps with a fast line segment detector (LSD) proposed by von Gioi et al. (2010). It is a robust line segment detection algorithm requiring no tuning of input parameters. The LSD partitions the image into line-support regions according to proximity and gradient angle. The regions are then approximated by rectangles having as the centre of mass the central point and uisng first inertia axis for rectangle orientation. The width and length of rectangles are chosen so that they cover all the points in the line-support region. All the line support regions are validated before the line segments are defined from the rectangles.

3. Matching of building outlines from hyperspectral images and DSM

Both object outline detection and line detection result in sets of lines in both data sets, HSI and DSM. A mismatch between the extracted line segments exists also because the accuracies of the georeferencing and orthorectification are imperfect. Both data sets (HSI and DSM) are represented as a projection onto the same 2D plane. Thus, a 2D affine transformation can be estimated to refine the coregistration. With our approach we aim to optimally reduce this mismatch by determining the transformation parameters between the data sets. For this purpose, the corresponding lines from both data sets need to be found first. We propose a method based on accumulating the possible line correspondences, which is described in Section 3.1. Using the corresponding line pairs, we estimate the transformation with the Gauss-Helmert model, which is described in Section 3.2.

3.1. Line assignment

In this subsection, we propose a method which can be applied also for data where the relative mismatch between the data sets amounts to 20-30 image pixels or even larger.

In our automatic approach for finding line correspondences we use a 3D accumulator. The 3D accumulator space is defined by a two dimensional translation in the x- and y-directions and a rotation. To fill the accumulator space we use one of the data sets as master image and move line segments from the second data set, i.e. slave image over the line segments from the master image. For each position we check how many lines correspond to each other and fill the accumulator with this number. At the same time we store the line correspondences for this position. We repeat this procedure by rotating line segments from the second data set by small angles in the range of few degrees. This algorithm results in a 3D accumulator space filled with line correspondences for each cell of the accumulator. The cell with maximal number of correspondences is selected and these line correspondences assigned to this cell are used for the transformation parameter calculation.

The correspondence between the lines for every cell of the accumulator is determined using statistical tests in consideration of the accuracy of the detected lines. We use homogeneous coordinates to formulate the problem. As shown in Heuel (2002) we calculate the distance vector $\mathbf{d}$ and test the hypothesis H_0

$$H_0 : \mathbf{d} = \mathrm{U}(\mathbf{e}_1)\mathbf{e}_2 = \mathrm{V}(\mathbf{e}_2)\mathbf{e}_1 = \mathbf{0} \tag{4}$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ are entities and U and V define the relation between $\mathbf{e}_1$ and $\mathbf{e}_2$. In this particular case, $\mathbf{e}_1$ and $\mathbf{e}_2$ are lines

and we investigate their incidence. Further on, we calculate the covariance matrix for the distance vector $\mathbf{d}$

$$\mathbf{\Sigma_{dd}} = \mathrm{U}(\mathbf{e}_1)\mathbf{\Sigma}_{e_2e_2}\mathrm{U}^{\mathrm{T}}(\mathbf{e}_2) + \mathrm{V}(\mathbf{e}_1)\mathbf{\Sigma}_{e_1e_1}\mathrm{V}^{\mathrm{T}}(\mathbf{e}_2), \tag{5}$$

where $\mathbf{\Sigma}_{e_1e_1}$ and $\mathbf{\Sigma}_{e_2e_2}$ are the covariance matrices for line $\mathbf{e}_1$ from the first data set and line $\mathbf{e}_2$ from the second data set, respectively. Then, we reject H_0 with a significance level α if

$$t = \mathbf{d}^{\mathrm{T}}\mathbf{\Sigma}_{\mathbf{dd}}^{-1}\mathbf{d} > \epsilon_H = \chi^2_{1-\alpha;n} \tag{6}$$

where ϵ_H is a critical value such that the probability satisfies equation $\mathrm{P}(t > \epsilon_H|H_0)$ defined by the hi-square distribution $\chi^2_{1-\alpha;\,n}$ with n degrees of freedom. Statistical tests are done for all the combinations between lines from the first data set with lines from the second data set, and are therefore computationally very expensive. To restrict possible line combinations and reduce computational efforts we generate H_0 only if the distance between the middle points of the line segments is smaller than a defined threshold. This threshold, similarly to the size of the accumulator, depends on the expected relative error of coregistration between the data sets.

3.2. Determination of transformation parameters

Selected correspondences are used for the determination of the transformation parameters. As mentioned in Section 3, HSI and DSM are represented as a projection onto the same 2D plane, therefore the coregistration refinement can be simplified with a 2D transformation. In general, a mapping between two planes for a 2D point $\mathbf{x}$ can be represented as

$$\mathbf{x'} = \mathbf{Hx} \tag{7}$$

where $\mathbf{x'}$ is transformed $\mathbf{x}$ and $\mathbf{H}$ is the transformation matrix. The vectors $\mathbf{x}$ and $\mathbf{x'}$ are represented using homogeneous coordinates. Basically, we want to map the line segments detected in the HSI to the height data. In theory, we can transform the end points of these line segments using (Eq. 7). Then, the corresponding end points of the line segments in the master image would be represented by $\mathbf{x'}$. However, in practice it is not possible to detect exactly identical line segments in both data sets. Usually, a part of the segment is missing in one data set, or the end points are not accurately determined. Therefore, we reformulate (Eq. 7) so that we still search for $\mathbf{H}$, but use lines as observations:

$$\mathbf{m} = \mathbf{H}^{-\mathrm{T}}\mathbf{l} \Rightarrow \mathbf{H}^{\mathrm{T}}\mathbf{m} = \mathbf{l} \tag{8}$$

where $\mathbf{m}$ represents a line indicated by a line segment from the master image and $\mathbf{l}$ represents a line indicated by a line segment from the slave image (see Hartley and Zisserman, 2004, pg. 36, Eq. 2.6). We assume that $\mathbf{m}$ and $\mathbf{l}$ represent the same line and use the condition for the identity relation as

$$\mathbf{H}^{\mathrm{T}}\mathbf{m} \times \mathbf{l} = 0. \tag{9}$$

The displacement between the data sets can be approximated by an affine transformation. Therefore, $\mathbf{H}$ is defined as

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ 0 & 0 & 1 \end{bmatrix}. \tag{10}$$

We calculate a line as the joining of two end points in the detected line segments $\mathbf{x}_1$ and $\mathbf{x}_2$

$$\mathbf{l} = \mathbf{x}_1 \times \mathbf{x}_2 \tag{11}$$

and normalize them spherically

$$\mathcal{N}(\mathbf{l}) = \frac{\mathbf{l}}{|\mathbf{l}|}. \tag{12}$$

Finally we estimate the optimal affine transformation parameters using the Gauss-Helmert model by minimizing the cost function

$$\Omega = \frac{1}{2}\hat{\mathbf{v}}_{\mathbf{b}}^{\mathrm{T}}\Sigma_{\mathbf{bb}}^{-1}\hat{\mathbf{v}}_{\mathbf{b}} + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{g}(\hat{\mathbf{b}}, \hat{\mathbf{p}}) + \boldsymbol{\nu}^{\mathrm{T}}\mathbf{c}(\hat{\mathbf{b}}) \tag{13}$$

subject to

$$\mathbf{g}(\hat{\mathbf{b}}, \hat{\mathbf{p}}) = \mathbf{H}^{\mathrm{T}}\mathbf{m} \times \mathbf{l} = 0 \tag{14}$$

and

$$\mathbf{c}(\hat{\mathbf{b}}) = |\mathbf{l}| - 1 = 0, \tag{15}$$

where $\hat{\mathbf{b}} = [\hat{\mathbf{l}}, \hat{\mathbf{m}}]^{\mathrm{T}}$ are observations, $\Sigma_{\mathbf{bb}} = diag[\Sigma_{\mathbf{ll}}, \Sigma_{\mathbf{mm}}]$ is their covariance matrix, $\hat{\mathbf{p}} = [h_1,\ h_2,\ h_3,\ h_4,\ h_5,\ h_6]^{\mathrm{T}}$ are the estimated parameters, $\hat{\mathbf{v}}_{\mathbf{b}} = [\hat{\mathbf{v}}_{\mathbf{l}}, \hat{\mathbf{v}}_{\mathbf{m}}]^{\mathrm{T}}$ are the corrections to line parameters of $(\mathbf{l}, \Sigma_{\mathbf{ll}})$ and $(\mathbf{m}, \Sigma_{\mathbf{mm}})$, and $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ are Lagrangian multipliers. The function $\mathbf{g}(\hat{\mathbf{b}}, \hat{\mathbf{p}})$ conditions the identity of the line pairs, while the function $\mathbf{c}(\hat{\mathbf{b}})$ constraints the spherical normalization for each line. We estimate the parameters $\hat{\mathbf{p}}$ in a similar way as described in Meidow et al. (2009), with the difference that we skip the restriction on the parameters. This restriction on the parameters is not needed for modeling the problem in the way presented in this paper.

4. Experiment

In this Section, we present our experiments carried out on real data. First, the used data sets are described (Subsec. 4.1), then the results of the proposed object-based coregistration refinement are presented and discussed (Subsec. 4.2). Finally, the evaluation of the estimated transformation parameters according to the ground truth, i.e. reference building footprints provided by the city of Munich, Gemrany is carried out (Subsec. 4.3).

4.1. Data description

The proposed object-based registration method is tested on two HSIs of different spatial and spectral resolution, and two DSMs of different spatial resolution (Tab. 1). All the data sets are acquired over the same residential area of the city of Munich, Germany. In addition, building footprints from the municipal of Munich, Germany are used as a reference for evaluation.

The two airborne hyperspectral data sets were acquired by the airborne HySpex and HyMap sensors and have a spatial resolution of 2 and 4 m respectively. The first data set was acquired by the HySpex hyperspectral system consisting of two sensors, VNIR and SWIR camera, providing two HSIs. For the tests we used only the VNIR image with 160 channels in the spectral range 0.4-1.0 μm, and 3.7 nm sampling interval. Four channels were removed before processing due to their high noise level. The second hyperspectral data set was acquired by the HyMap sensor and has 125 channels in the spectral range 0.4-2.5 μm. 24 bands were removed due to the noise level, e.g. noise caused by atmospheric absorption regions. Both HSIs are corrected to at-sensor radiance, but no atmospheric correction was applied.

Type of data set	Image size [pix]	GSD [m]
Hyperspectral		
HySpex	279 x 370	2
HyMap	141 x 185	4
DSM		
LiDAR	579 x 759	1
3K DSM	279 x 370	2

Table 1: List of tested data sets with ground sampling distance (GSD) and image size. The size of the test area is approximately 0.76 x 0.58 km. The number of extracted and matched outlines and lines from each of the data sets is listed in Tab. 2.

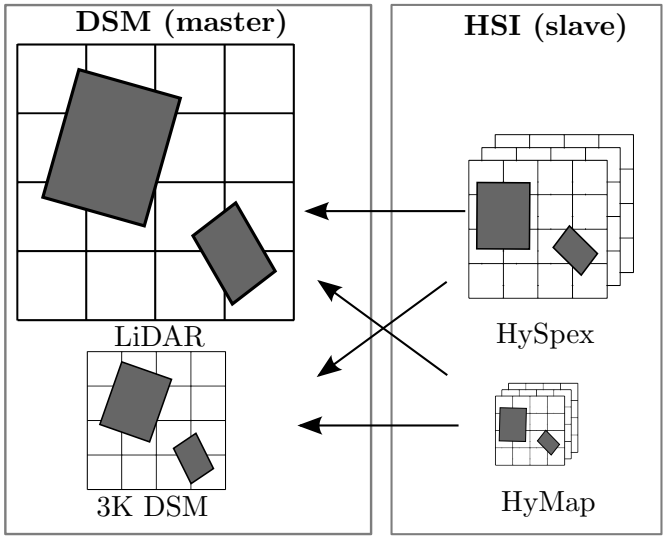


Figure 4: Registration between master DSM and slave HSI. Eight registrations are carried out: each DSM is registered with each HSI, once using detected objects, and once using detected line segments. The relative size between the grids corresponds to the spatial resolution of the data sets.

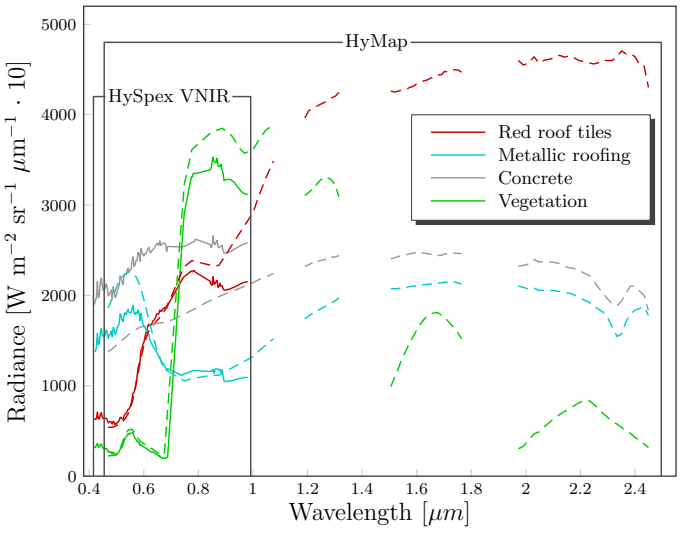


Figure 5: Spectral signatures of four materials. Spectra collected from HyMap (dashed lines) and HySpex (solid lines) for three roofing materials and vegetation. THe concrete spectral signature is not pure, due to the HSI spatial resolution and to the presence of smaller objects on concrete roofs. The spectral domain of the HySpex VNIR and the HyMap sensor is marked with boxes. The HSI data sets are radiometrically corrected to at-sensor radiance.

The two DSMs used as master images in the coregistration process are the LiDAR DSM and the 3K DSM. The former, with a spatial resolution of 1 m, is resampled from a LiDAR point cloud with an average density of 1.69 points/m^2. The latter, with a spatial resolution of 2 m, is computed from multi-view optical images using semi-global matching. The optical images were acquired by the 3K-camera system, consisting of three non-metric cameras, of which one is nadir looking and two two are oblique sidewards looking. The airborne 3K system was developed at DLR IMF (Kurz et al., 2012).

4.2. Results and discussion

The coregistration refinement of the HSI (slave image) to the DSM (master image) is carried out for all four possible combinations between the data sets as plotted in Fig. 4. Furthermore, the coregistration for each pair consisting of a DSM and a HSI is carried out twice: once on the basis of detected objects and once using detected line cues. So, altogether eight coregistrations between the data sets of different spatial (1-4 m) and spectral resolution are calculated. In addition, another 16 coregistrations between reference building footprints from the city of Munich, Germany and all the detected building outlines are computed for evaluation purposes (Subsec. 4.3).

Both HSI are unmixed with the NNLS method, using spectral signatures collected directly from the image. Hence, four reference spectra are collected from the HyMap image and four reference spectra from the HySpex image. Results are four material maps for HyMap and HySpex, one for vegetation and three for roofing materials. For each reference spectral signature, ten spectra are collected from the HSI and averaged to suppress possible noise (Fig. 5). The reference spectral signatures from HyMap (dashed lines) and HySpex (solid lines) HSIs are plotted in Fig. 5. The boxes represent the spectral domain of each sensor, with missing values indicating the removed noisy channels (Subsec. 4.1). The reference spectra for vegetation, concrete and red roofing tiles correspond to the expected spectral signatures. However, the collected image spectra for metallic roofing and concrete are most probably not pure pixels, because of smaller objects on the roofs and the low spatial resolution

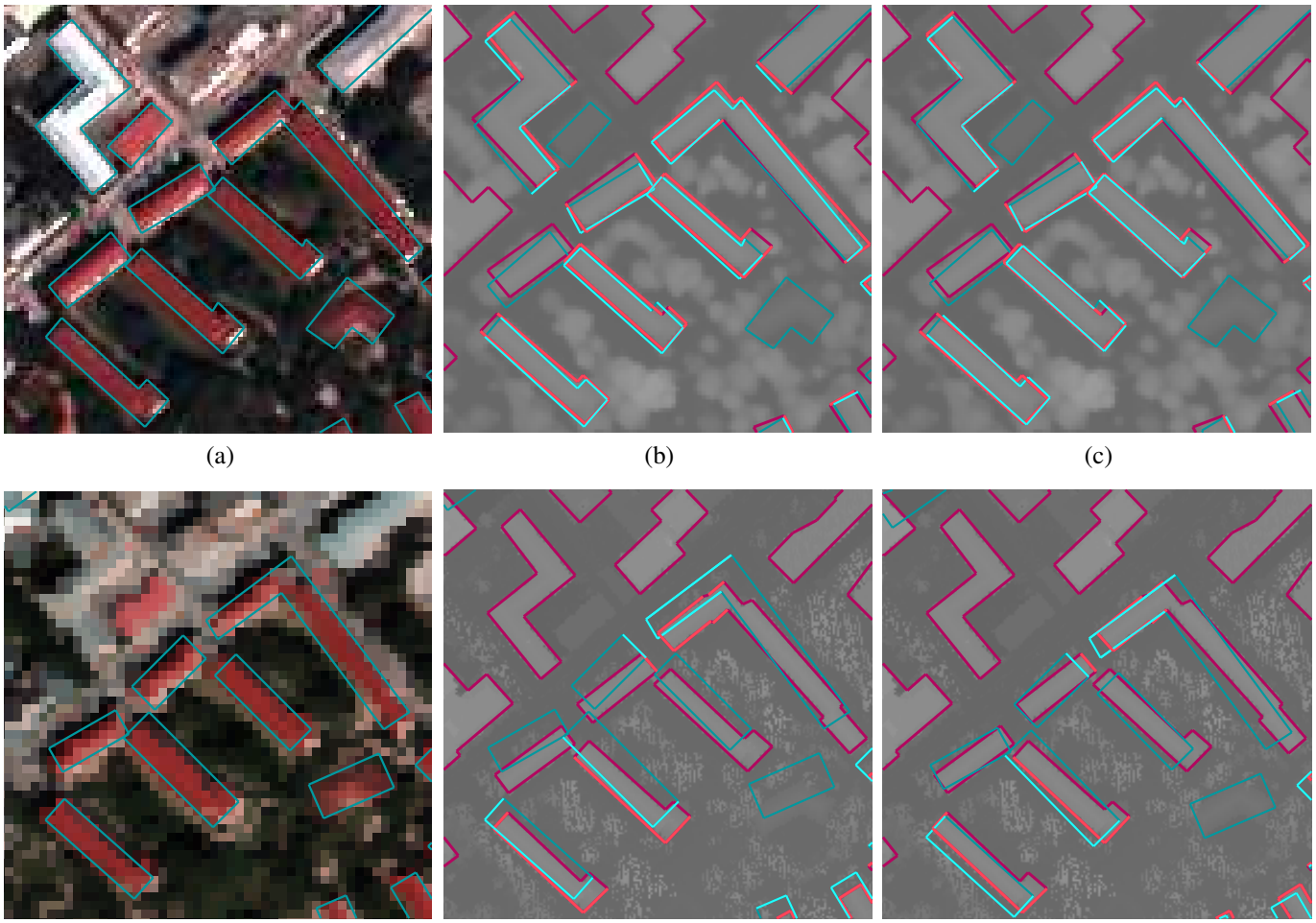


Figure 6: Detected building outlines from two HSI (a, d) and two DSM (b, e), before (b, e) and after coregistration (c, f). True colour composites of a of HySpex (a; 100x100 pixels) and a HyMap (d; 50x50 pixels) subsets with overlaid detected building outlines (cyan). Two colors are used for extracted building outlines: cyan and magenta for outlines extracted from the slave HSI, and magenta for the ones from the master DSM, respectively. If correspondence between two line segments in master and slave image is found, the color of this line segment is highlighted, i.e. bright cyan and bright magenta.

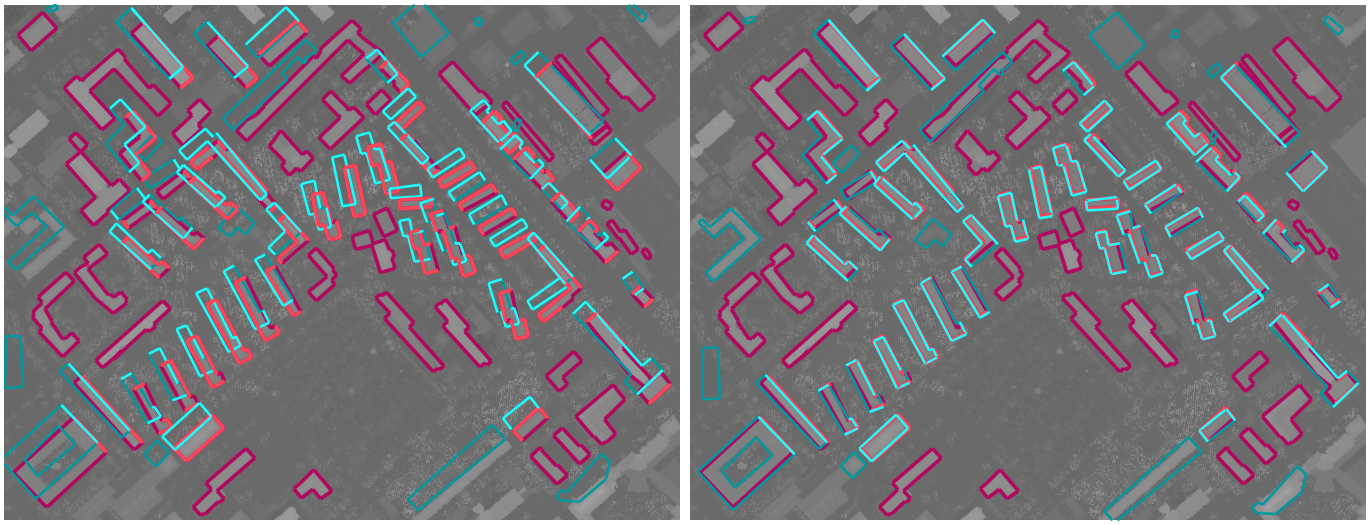
of the HSI. Nevertheless, we keep all material maps for building outline extraction. Furthermore, if HSIs are radiometrically and atmospherically corrected, the reference spectra could be collected from only one HSI or taken from a spectral library. However, no atmospheric correction was applied to HyMap and HySpex HSIs, thus reference spectra are collected from each image separately.

HySpex reference spectral signatures exhibit some sharp edges, indicating not corrected systematic artefacts in the image. However, noise is present in the whole HySpex image and also in the reference spectral signatures collected from the same image. Thus, it does not significantly influence the NNLS unmixing. On the contrary, using spectral library spectra or spectra from HyMap would influence unmixing result of the HSI with systematic errors.

The proposed building outline detection (Subsec. 2.2) provides a large enough number of approximate building outlines for line assignment and coregistration between all pairs of HSI and DSM (Tab. 2). Two thresholds are set to obtain the initial building mask from the HSIs and the DSMs. The

no. outlines (*no. lines*)	**LiDAR** 486 (543)	**3K DSM** 409 (348)	**Reference footprint** 1534
HyMap 94 (88)	51 (30)	45 (22)	66 (28)
HySpex 278 (255)	187 (95)	147 (70)	208 (92)
Reference footprints 1534	524 (541)	239 (175)	-

Table 2: Number of detected and matched lines and outlines. Under each data set (**bold**) is a number of line segments of detected building outlines (*no. outlines*) and in brackets a number of detected lines (*no. line*). Numbers of matched lines and outlines between pairs of data sets are in the fields of the table. Each outline consists of several line segments, so the number of outlines in this Table is counted as the total number of line segments of all the building outlines.



(a) (b)

Figure 7: Detected building outlines in LiDAR DSM and HySpex HSI, before (a) and after coregistration (b) overlaid on LiDAR DSM of 579x759 pixels. Outlines detected in the LiDAR DSM are in magenta, while and outlines detected in the HySpex HSI are in cyan. Matching lines are highlighted (bright magenta for HSI, bright cyan for DSM), analogue to the description under Fig. 6.

threshold for material maps from HyMap and HySpex is set to $\mathbf{BM}_{HSI,i} > 0.7$, $i = 1,\ldots,3$, for three roofing materials, and the threshold for normalized 3K DSM and LiDAR DSM is set to $\mathbf{BM}_{DSM} > 2$ m. The threshold for the initial building mask $\mathbf{BM}_{HSI,i} > 0.7$ means that at least 70% of pixel's area is covered by the roofing material. This threshold is based on heuristics and does not change building outline estimation when it is in a range 0.7 ± 0.1. The abundance maps and the normalized DSMs are used to estimate the building outlines, whereas the thresholded material maps and thresholded normalized DSM are used only for an initial guess for building regions (Subsec. 2.2.1). For a majority of the building outlines the automatically selected building model fits the data, and is detailed enough to represent different building shapes. In data sets with different spatial resolutions a building can be estimated with a model of different complexity levels. For example, in Fig. 6 the same observed area with detected building outlines in HySpex (a) and HyMap (d) is shown with GSD of 2 m and 4 m, respectively. Some of the same buildings in the scene are approximated with different building outline models in different datatets. Thus, for the building outlines in the HyMap with lower GSD, simpler building models, e.g. rectangle, are chosen than in the HySpex with higher GSD, e.g. L-shape. However, in all four used data sets some buildings only partialy present in the image are badly estimated or not estimated at all (Fig. 7). Some of the buildings in HSI are not extracted, because only three roof materials were considered. This can be observed by comparing the number of extracted building line segments in 3K DSM (409) and HySpex (278), with both sensors having the same GSD (Tab. 2). The average runtime for the given data sets is about 1.1 s per building (using not optimized prototype software on a computer with a 2.4GHz dual core processor). The runtime approximately linearly depends on the number of rectangles required for representing a building outline. To increase the completeness of building extraction, all building materials in the given scene should be collected. Next to the detected outlines, also line cues are extracted with LSD algorithm (Subsec. 2.2.2). We use pre-defined parameters and a suggested Gaussian smoothing of the image with standard deviation $\sigma = 0.8$.

Correspondences between the data sets were found using the accumulator approach. The movement of the reference point was set to ±25 pixels in each direction and small rotation angles to ±0.5°. The significance level for the statistical test (Eq. 6) was set to $\alpha = 0.08$. Then, the number of selected correspondences were counted and presented in Tab. 2. In this table, the number of line segments extracted during the outline detection is presented below the data description. The number of line cues detected with LSD (Subsubsec. 2.2.2) is specified in brackets. The runtime of the accumulator for matching, e.g. LiDAR outlines and HyMap outlines, is about 16 s (using not optimized prototype software on a computer with a 3.1GHz dual core processor). The runtime is approximately lineary dependant on the number of cells in the accumulator and on the number of tested hypothesis.

A comparison of the number of matched outlines and lines is made for all the data sets (Tab. 2), including the number of reference building outlines, i.e. footprints. For all tested data sets (Fig. 4) more assignments are found between the extracted building outlines than between extracted line cues with LSD (Subsubsec. 2.2.2). Concerning the spatial resolution of the data, images with higher spatial resolution include more details, and so more outline segments are extracted. Consequently, more matched line segments between the data sets with higher GSD are found.

The initial registration accuracy of the LiDAR DSM and both

HSIs was reduced to demonstrate the effectiveness of the accumulator for matching (Fig. 7, Fig. 6, e, f). The initial registration for the 3K DSM and HySpex is about one pixel (Fig. 6). For all tested data sets (Fig. 4) the estimated transformation matrix $\mathbf{H}_{\text{est}}$, i.e. the affine parameters $h_1 - h_6$ with standard deviations (Eq. 10) are estimated. The covariance matrices for the line parameters used in the adjustment were calculated by error propagation from the covariance matrices of the end points.

Before estimating the transformation parameters, the observations were conditioned similar as shown in e.g. Heuel (2002), Zeng et al. (2008). Using line parameters as observations the conditioning of the observed parameters is more important as for points (Zeng et al., 2008). While using the covariance matrices for the adjustment, bad conditioning leads to numerical problems and disables the adjustment procedure.

As expected, the highest impact on the transformation are due to the translation parameters (h_3, h_6). The other parameters vary insignificantly (in most cases < 0.005) from the standard value $(h_1 = 1, h_2 = 0, h_4 = 0, h_5 = 1)$. The standard deviation for h_1, h_2, h_4, h_5 ranges from 0.0001 to 0.002, while for h_3, h_6 it ranges from 0.065 meters (LiDAR-HySpex) to 0.186 meters (LiDAR-HyMap).

4.3. Evaluation

The evaluation of the proposed method can be done by investigating the computed inner accuracies of the estimated transformation matrix $\mathbf{H}_{\text{est}}$, i.e. standard deviations of elements. Such an assessment however, does not provide the comparison to the true transformation which has to be carried out to coregister both data sets. Hence, a reference transformation has to be determined and the absolute difference between the estimated $\mathbf{H}_{\text{est}}$ transformation parameters and the reference $\mathbf{H}_{\text{ref}}$ parameters has to be computed. Additionally, correctness and quality rate of extracted building outlines before and after coregistration refinement is computed.

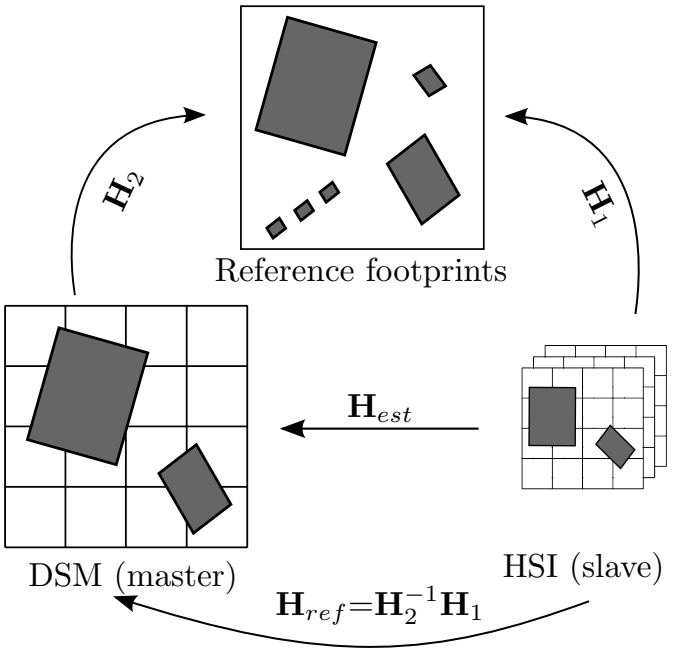


Figure 8: Evaluation of the computed projection matrix $\mathbf{H}$. The projection matrix between DSM (master) and HSI (slave) $\mathbf{H}$ is compared to the $\mathbf{H}_{\text{ref}} = \mathbf{H}_2^{-1}\mathbf{H}_1$. The projection matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ project the objects respectively from DSM and HSI to the auxiliary coordinate system of ground truth building outlines.

A manually selection of tie points from in the HSI is insufficient for computing a reference transformation matrix (Avbelj et al., 2013). The position of a manually set tie point in a HSI strongly depends on the contrast and selection of the channels from the HSI. Our method for coregistration refinement is based on correspondences between the object outlines from the slave and the master images, but extracted outlines from both images include some inaccuracies. So, to compute a reference transformation matrix $\mathbf{H}_{\text{ref}}$, we use the auxiliary data (building footprints) with much higher accuracy than the extracted building outlines (Fig. 8). For the reference building footprints the data from the city of Munich, Germany are used.

If the affine matrix $\mathbf{H}_1$ projects points from the slave coordinate system to the auxiliary coordinate system and the affine matrix $\mathbf{H}_2$ projects points from master coordinate system to the same auxiliary coordinate system, then $\mathbf{H}_{\text{ref}} = \mathbf{H}_2^{-1}\mathbf{H}_1$ should be equal to the projection as observed directly from the slave to the master coordinate system $\mathbf{H}_{\text{est}}$ (Fig. 8). The reference footprints in the auxiliary coordinate system are assumed to be error free, nevertheless the $\mathbf{H}_1$ and the $\mathbf{H}_2$ still include the inaccuracies of the outline (line cue) extraction from the master or the slave image. These inaccuracies are also result from the different GSDs of the images, where the GDS limits the smallest extractable detail and the accuracy of the extracted line segments. Yet, matching extracted outlines to the reference footprints excludes the uncertainty of the extracted line segments in the auxiliary coordinate system. Moreover, some inaccurate matches between the reference footprints and extracted outlines (line clues) are avoided, because the errors in outline extraction are only expected in the slave and the master images. Thus, the $\mathbf{H}_{\text{ref}} = \mathbf{H}_2^{-1}\mathbf{H}_1$ is more reliable and accurate then $\mathbf{H}_{\text{est}}$, where both outlines from slave and master image introduce inaccuracies into the matching and estimation of transformation parameters.

In Fig. 9 the comparison between the elements of reference $\mathbf{H}_{\text{ref}}$ and estimated $\mathbf{H}_{\text{est}}$ transformation matrices is represented. Horizontal columns represent the absolute difference of six elements of reference and estimated projection matrix $|\mathbf{H}_{\text{ref}} - \mathbf{H}_{\text{est}}|$ for all the data sets, and for extracted outlines (dark grey) and lines (light grey). Parameters $h_3 = dx$ and $h_6 = dy$ are the translations in x and y direction in *meters*, the other h affine parameters are dimensionless quantities. The error bars on each horizontal column are the computed standard deviations of $\mathbf{H}_{\text{est}}$. The differences between estimated and reference transformation matrix are larger for the data sets with lower GSD, i.e. 3K DEM and HyMap. The estimated standard deviation of the parameters, i.e. inner accuracies of estimation are in most cases too optimistic, compared to the difference between reference and estimated transformation parameters.

Correctness and quality rate of extracted building outlines using initial georeferencing and after coregistration refinement for all pairs of coregistrated images are given in Tab. 3 and Tab. 4,

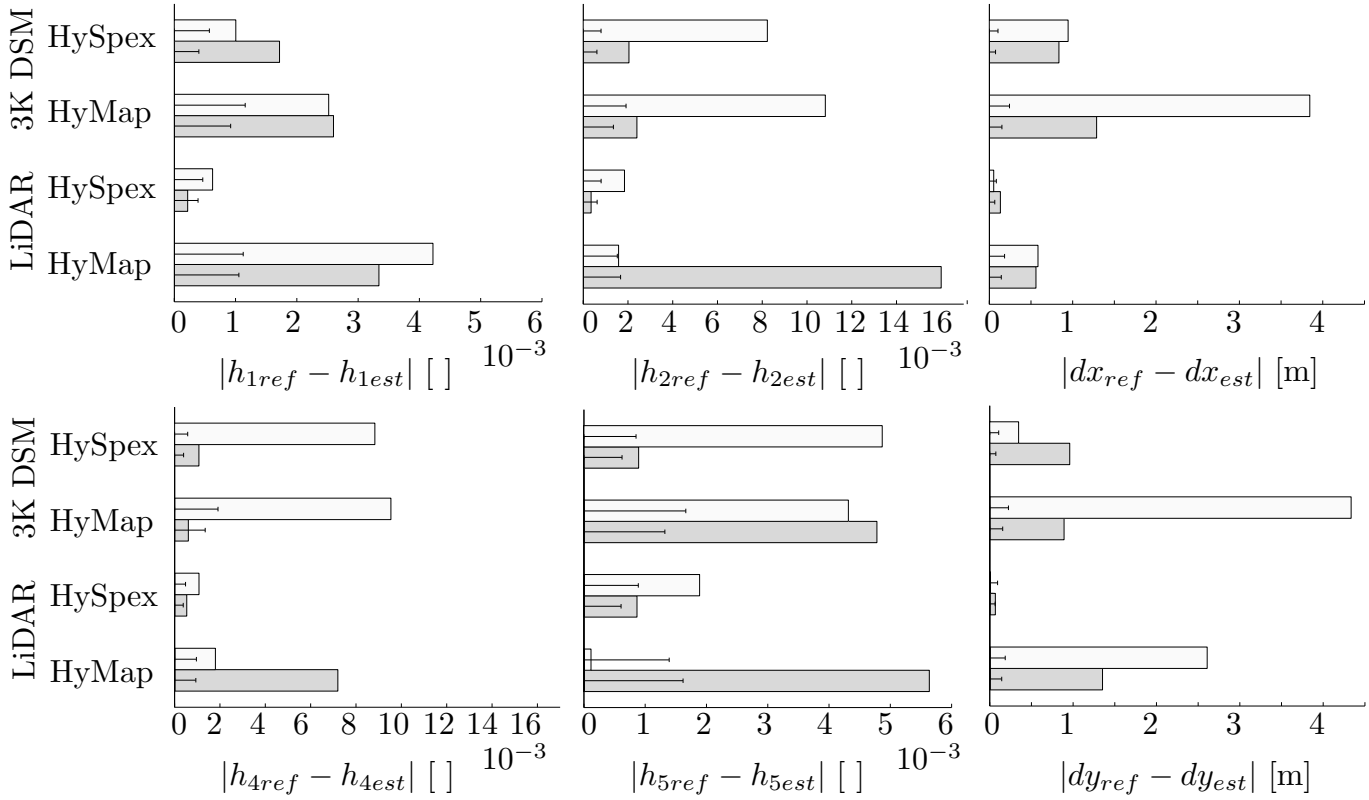


Figure 9: Evaluation of the estimated projection matrix $\mathbf{H}_{\text{est}}$ for all data sets. The estimated projection matrix $\mathbf{H}_{\text{est}}$ between DSM (master) and HSI (slave) is compared to the reference projection matrix $\mathbf{H}_{\text{ref}} = \mathbf{H}_2^{-1}\mathbf{H}_1$. For each of the six elements in the projection matrix (eq. 10), the absolute difference between the reference and the estimated value is computed (horizontal axis). This difference is represented with grey horizontal bars on the six charts. Dark grey bars represent the values for matched outlines, and light grey for matched line cues (vertical axis). The error bars (horizontal lines on every bar) represent the inner accuracies of the $\mathbf{H}_{\text{est}}$ and are computed as the standard deviations of the six elements of the $\mathbf{H}_{\text{est}}$. The absolute differences of the translations are in *meters*, the other absolute differences are dimensionless quantities.

Correctness $\mathbf{H}_{\text{outlines}}$ ($\mathbf{H}_{\text{lines}}$)		**LiDAR**	**3K DSM**
HyMap	*Initial*	0.54	0.62
	Refined	0.63 (0.64)	0.77 (0.70)
HySpex	*Initial*	0.31	0.77
	Refined	0.64 (0.58)	0.79 (0.79)

Table 3: Correctness of extracted building outlines using initial georeferencing (*Initial*), and after the coregistration refinement (*Refined*). Correctness is computed as the ratio between the overlap area of extracted building footprints and the total area of all extracted building footprints in slave (DSM) image. The extracted building outlines are transformed using the estimated transformation matrix ($\mathbf{H}_{\text{est}}$), where two different $\mathbf{H}_{\text{est}}$ are used for each pair of master-slave dataset. First, $\mathbf{H}_{\text{est}}$ is estimated from the extracted building outlines ($\mathbf{H}_{\text{outlines}}$), and second from the extracted line cues ($\mathbf{H}_{\text{lines}}$). The rows denoted *Refined* have two values of correctness for each pair of master-slave images, i.e. correctness of extracted building footprints transformed by $\mathbf{H}_{\text{outlines}}$ (value without brackets) and by $\mathbf{H}_{\text{lines}}$ (value in brackets). The rows denoted *Initial* have a correctness value using initial georeferencing.

Quality rate $\mathbf{H}_{\text{outlines}}$ ($\mathbf{H}_{\text{lines}}$)		**LiDAR**	**3K DSM**
HyMap	*Initial*	0.20	0.18
	Refined	0.24 (0.24)	0.23 (0.21)
HySpex	*Initial*	0.16	0.40
	Refined	0.41 (0.36)	0.42 (0.42)

Table 4: Quality rate of extracted building outlines using initial georeferencing (*Initial*), and after the coregistration refinement (*Refined*). Quality rate is computed as the ratio between the overlap area of extracted building footprints and the union of all areas of extracted and reference building footprints. The are notation is analogue to the description under Tab. 3.

respectively. Correctness is computed as the ratio between the overlapping area of extracted building footprints and the total area of all extracted building outlines in the slave image. Quality rate is computed as the ratio between the overlapping area of extracted building footprints and the union of the area of all building outlines. The quality rate, in contrast to correctness, accounts for the areas of building polygons that are not extracted. Both, correctness and quality rate quantify with a single value the quality of the building outline extraction and the coregistration refinement.

According to the results shown in Fig. 9 and Tab. 3 using

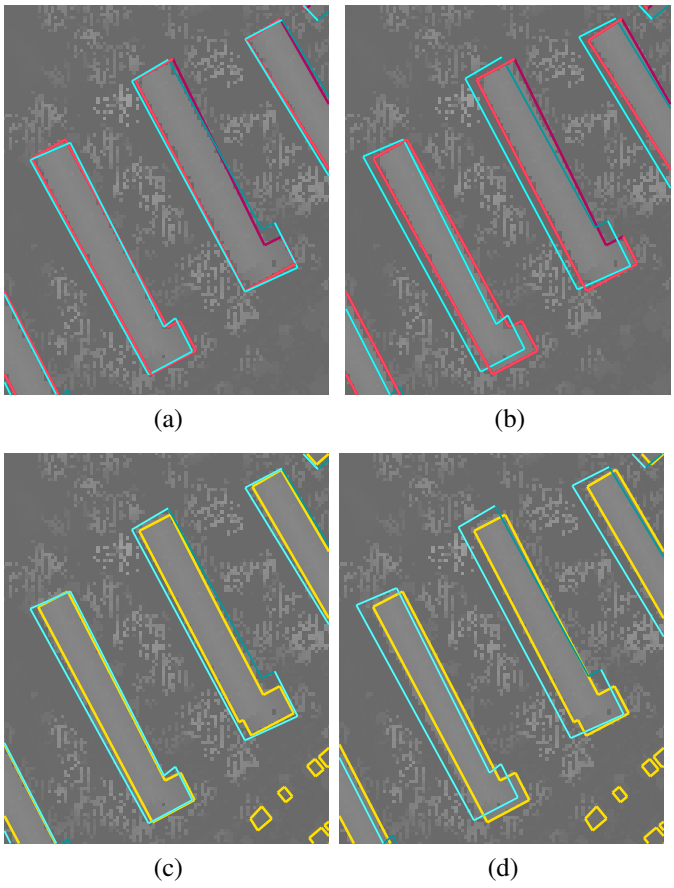


Figure 10: Building outlines detected in the slave HySpex HSI (cyan) after coregistration, projected using a transformation matrix on the basis of matched outlines (a, c) and lines (b, d) from master LiDAR DSM (magenta). Subfigures c and d show the reference building outlines (yellow) and the detected building outlines in HySpex (cyan). All outlines are overlaid on the detail of 120x100 pixels from the LiDAR DSM. Matching lines are highlighted analogue to the description under Fig. 6.

outlines of objects instead of the image cues is advantageous for coregistration for all the tested data sets, except for HyMap and LiDAR DSM. This conclusion is based on jointly observing all values of the transformation matrix $|\mathbf{H}_{\text{ref}} - \mathbf{H}_{\text{est}}|$ and correctness values of extracted building outlines. Comparison of the standard deviations of the $\mathbf{H}_{\text{est}}$ (error bars, Fig. 9) to the $|\mathbf{H}_{\text{ref}} - \mathbf{H}_{\text{est}}|$ serves mainly as a plausibility check for the proposed evaluation technique. Building outlines in HySpex DSM projected with the transformation matrix computed from building outlines (Fig. 10, a), fit better to the image data, than projected with the transformation matrix computed from line cues (Fig. 10, b). What is more, also the detected building outlines in HySpex (cyan lines in Fig. 10) fit better to reference building outlines (magenta lines in Fig. 10, c, d, Tab. 3), when projected with the transformation matrix computed from building outlines.

5. Conclusion and future work

We have proposed an object-based coregistration refinement method for HSI and DSM using homogeneous coordinates for line assignment and transformation parameter estimation. We showed that the proposed method is applicable on multimodal and multi-resolution data sets with different characteristics. We have focused the coregistration on building outlines, however other objects in the scene could be used if they can be extracted on the basis of material properties and height information.

Detecting outlines of the building objects in HSI and DSM implicitly incorporates topology relations and metric restrictions between the subsequent line segments of each building outline. Estimating building outlines instead of line segments provide a larger amount of assigned lines between the images and is thus more reliable and also more robust for estimation of transformation. The results indicate that additional topology and metric relations of the outline of the observed objects is especially important for images with lower spatial resolution. The method has its limitation when buildings to be extracted have all sides smaller than three pixels. This finding is consistent with the discussion of Blaschke (2010) about improvements of object-based over per-pixel image analysis.

We have introduced geometric accuracy of detected line segments on building outlines for optimal transformation parameter estimation. The used stochastic model provides results. So, in the future investigations on the influence of line accuracies on transformation parameters should be carried out. Moreover, for fully operational coregistration refinement processing systems, the computational time, robustness to input parameters, and economical aspects should be investigated and improved.

The correctness and quality rate values of extracted building outlines after coregistration refinement are improved in comparison to the initial georeferencing. We have additionally proposed an approach for the evaluation of coregistration refinements, as the computed standard deviations of estimated transformation parameters are not sufficient. The reference transformation parameters are computed by using a much better vector description of the objects under consideration, in our case reference building outlines.

Acknowledgement

The authors would like to thank the city of Munich, Department of Environment and Health (Referat für Gesundheit und Umwelt der Stadt München) for providing the reference building outlines (footprints) used as ground truth.

References

Avbelj, J., 2012. Spectral information retrieval for sub-pixel building edge detection. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences I (7), 61–66. doi:10.5194/isprsannals-I-7-61-2012.

Avbelj, J., Iwaszczuk, D., Müller, R., Reinartz, P., Stilla, U., 2013. Line-based registration of DSM and hyperspectral images. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL (1/W1), 13–18. doi:10.5194/isprsarchives-XL-1-W1-13-2013.

Awrangjeb, M., Ravanbakhsh, M., Fraser, C.S., 2010. Automatic detection of residential buildings using LIDAR data and multispectral imagery. ISPRS Journal of Photogrammetry and Remote Sensing 65 (5), 457–467. doi:10.1016/j.isprsjprs.2010.06.001.

Blaschke, T., 2010. Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing 65 (1), 2–16. doi:10.1016/j.isprsjprs.2009.06.004.

Brédif, M., Tournaire, O., Vallet, B., Champion, N., 2013. Extracting polygonal building footprints from digital surface models: A fully-automatic global optimization framework. ISPRS Journal of Photogrammetry and Remote Sensing 77 (1), 57–65. doi:10.1016/j.isprsjprs.2012.11.007.

Brook, A., Ben-Dor, E., 2011. Automatic registration of airborne and spaceborne images by topology map matching with SURF processor algorithm. Remote Sensing 3 (1), 65–82. doi:10.3390/rs3010065.

Brook, A., Ben-Dor, E., Richter, R., 2010. Fusion of hyperspectral images and LiDAR data for civil engineering structure monitoring. In: 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Reykjavik, Iceland, 14–16 June, pp. 1–5. doi:10.1109/WHISPERS.2010.5594872.

Chang, C.I., 2000. An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis. IEEE Transactions on Information Theory 46 (5), 1927–1932. doi:10.1109/18.857802.

Debevec, P., Taylor, C., Malik, J., 1996. Modelling and rendering architecture from photographs: A hybrid geometry- and image-based approach. In: Proc. of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, pp. 11–20. doi:10.1145/237170.237191.

Fischler, M., Bolles, R., 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM 24 (6), 381–395.

GEO/CEOSS, 2010. A quality assurance framework for earth observation, version 4.0. URL: http://qa4eo.org/docs/QA4EO_Principles_v4.0. (Accessed 23 Mar. 2014).

Gerke, M., Straub, B.M., Koch, A., 2001. Automatic detection of buildings and trees from aerial imagery using different levels of abstraction. In: Publikationen der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung 10 (1), pp. 273–280.

von Gioi, R., Jakubowicz, J., Morel, J.M., Randall, G., 2010. LSD: a fast line segment detector with a false detection control. IEEE Transactions on Pattern Analysis and Machine Intelligence 32 (4), 722 –732. doi:10.1109/TPAMI.2008.300.

Gröger, G., Kolbe, T.H., Nagel, C., Häfele, K.H., 2012. OGC city geography markup language (CityGML) encoding standard, version 2.0.0., Open Geospatial Consortium.

Gröger, G., Plümer, L., 2012. CityGML interoperable semantic 3D city models. ISPRS Journal of Photogrammetry and Remote Sensing 71 (1), 12–33. doi:10.1016/j.isprsjprs.2012.04.004.

Habbecke, M., Kobbelt, L., 2010. Automatic registration of oblique aerial images with cadastral maps. In: ECCV Workshop on Reconstruction and Modeling of Large Scale 3D Virtual Environments 6554 (1), pp. 253–266.

Hartley, R.I., Zisserman, A., 2004. Multiple View Geometry in Computer Vision. Second ed., Cambridge University Press, ISBN: 0521540518.

Heuel, S., 2002. Statistical Reasoning in Uncertain Projective Geometry for Polyhedral Object Reconstruction. Ph.D. thesis. Institute of Photogrammetry, University of Bonn, Germany.

INSPIRE TWG BU, 2013. Data specification on buildings: Draft technical guidelines, version: v3.0rc3, identifier: D2.8.III.2_ v2.9. Thematic working group buildings. URL: http://inspire.jrc.ec.europa.eu/index.cfm/pageid/2 (Accessed 24 Sep. 2013).

Iwaszczuk, D., Hoegner, L., Schmitt, M., Stilla, U., 2012. Line based matching of uncertain 3D building models with IR image sequences for precise texture extraction. Photogrammetrie, Fernerkundung, Geoinformation (PFG) 2012 (5), 511–521. doi:10.1127/1432-8364/2012/0135.

Keshava, N., 2003. A survey of spectral unmixing algorithms. Lincoln Laboratory Journal 14 (1), 55–78.

Kruse, F., Lefkoff, A., Boardman, J., Heidebrecht, K., Shapiro, A., Barloon, P., Goetz, A., 1993. The spectral image processing system (SIPS) interactive visualization and analysis of imaging spectrometer data. Remote Sensing of Environment 44 (2-3), 145–163. doi:10.1016/0034-4257(93)90013-N.

Kurz, F., Türmer, S., Meynberg, O., Rosenbaum, D., Runge, H., Reinartz, P., Leitloff, J., 2012. Low-cost optical camera systems for real-time mapping applications. Photogrammetrie, Fernerkundung, Geoinformation (PFG) 2012 (2), 159–176. doi:10.1127/1432-8364/2012/0109.

Lafarge, F., Descombes, X., Zerubia, J., Pierrotdeseilligny, M., 2008. Automatic building extraction from DEMs using an object approach and application to the 3D-city modeling. ISPRS Journal of Photogrammetry and Remote Sensing 63 (3), 365–381. doi:10.1016/j.isprsjprs.2007.09.003.

Maas, H.G., Vosselman, G., 1999. Two algorithms for extracting building models from raw laser altimetry data. ISPRS Journal of Photogrammetry and Remote Sensing 54 (2-3), 153–163. doi:10.1016/S0924-2716(99)00004-0.

Meidow, J., Beder, C., Förstner, W., 2009. Reasoning with uncertain points, straight lines, and straight line segments in 2D. ISPRS Journal of Photogrammetry and Remote Sensing 64 (2), 125–139. doi:10.1016/j.isprsjprs.2008.09.013.

Ok, A., Wegner, J., Heipke, C., Rottensteiner, F., Sörgel, U., Toprakc, V., 2012. Matching of straight line segments from aerial stereo images of urban areas. ISPRS Journal of Photogrammetry and Remote Sensing 74 (1), 133–152. doi:10.1016/j.isprsjprs.2012.09.003.

Pluim, J.P.W., Maintz, J.B.A., Viergever, M.A., 2003. Mutual-information-based registration of medical images: A survey. IEEE Transactions on Medical Imaging 22 (8), 986–1004. doi:10.1109/TMI.2003.815867.

Rottensteiner, F., Sohn, G., Gerke, M., Wegner, J.D., Breitkopf, U., Jung, J., 2014. Results of the ISPRS benchmark on urban object detection and 3D building reconstruction. ISPRS Journal of Photogrammetry and Remote Sensing in press. doi:10.1016/j.isprsjprs.2013.10.004.

Schaepman-Strub, G., Schaepman, M.E., Painter, T.H., Dangel, S., Martonchik, J.V., 2006. Reflectance quantities in optical remote sensing-definitions and case studies. Remote Sensing of Environment 103 (1), 27–42. doi:10.1016/j.rse.2006.03.002.

Schenk, T., 2004. From point-based to feature-based aerial triangulation. ISPRS Journal of Photogrammetry and Remote Sensing 58 (5-6), 315–329. doi:10.1016/j.isprsjprs.2004.02.003.

Schmid, C., Zisserman, A., 2000. The geometry and matching of lines and curves over multiple views. International Journal of Computer Vision 40 (3), 199–233.

Sohn, G., Dowman, I., 2007. Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. ISPRS Journal of Photogrammetry and Remote Sensing 62 (1), 43–63. doi:10.1016/j.isprsjprs.2007.01.001.

Stilla, U., 1995. Map-aided structural analysis of aerial images. ISPRS Journal of Photogrammetry and Remote Sensing 50 (4), 3–10. doi:10.1016/0924-2716(95)98232-0.

Suri, S., Reinartz, P., 2010. Mutual-information-based registration of TerraSAR-X and ikonos imagery in urban areas. IEEE Transactions on Geoscience and Remote Sensing 48 (2), 939–949. doi:10.1109/TGRS.2009.2034842.

Zeng, H., Deng, X., Hu, Z., 2008. A new normalized method on line-based homography estimation. Pattern Recognition Letters 29 (9), 1236–1244.

Zitova, B., Flusser, J., 2003. Image registration methods: A survey. Image and vision computing 21 (11), 977–1000.